Collider bias in trauma comparative effectiveness research: The stratification blues for systematic reviews

Deborah J. del Junco [a,*], Eileen M. Bulger [b], Erin E. Fox [a], John B. Holcomb [a], Karen J. Brasel [c], David B. Hoyt [d], James J. Grady [e], Sarah Duran [a], Patricia Klotz [b], Michael A. Dubick [f], Charles E. Wade [a] On Behalf of the ROC Investigators

[a] University of Texas Health Science Center, Department of Surgery, Houston, TX, United States
[b] University of Washington, Department of Surgery, Seattle, WA, United States
[c] Oregon Health & Science University, Department of Surgery, Portland, OR, United States
[d] American College of Surgeons, Chicago, IL, United States
[e] University of Connecticut Health Center, Institute for Clinical and Translational Science, Farmington, CT, United States
[f] U.S. Army Institute of Surgical Research, San Antonio, TX, United States

ARTICLE INFO

ABSTRACT

Background: Collider bias, or stratifying data by a covariate consequence rather than cause (confounder) of treatment and outcome, plagues randomised and observational trauma research. Of the seven trials of prehospital hypertonic saline in dextran (HSD) that have been evaluated in systematic reviews, none found an overall between-group difference in survival, but four reported significant subgroup effects. We hypothesised that an avoidable type of collider bias often introduced inadvertently into trauma comparative effectiveness research could explain the incongruous findings.
Methods: The two most recent HSD trials, a single-site pilot and a multi-site pivotal study, provided data for a secondary analysis to more closely examine the potential for collider bias. The two trials had followed the a priori statistical analysis plan to subgroup patients by a post-randomisation covariate and well-established surrogate for bleeding severity, massive transfusion (MT), $\geq$10 unit of red blood cells within 24 h of admission. Despite favourable HSD effects in the MT subgroup, opposite effects in the non-transfused subgroup halted the pivotal trial early. In addition to analyzing the data from the two trials, we constructed causal diagrams and performed a meta-analysis of the results from all seven trials to assess the extent to which collider bias could explain null overall effects with subgroup heterogeneity.
Results: As in previous trials, HSD induced significantly greater increases in systolic blood pressure (SBP) from prehospital to admission than control crystalloid ($p = 0.003$). Proportionately more HSD than control decedents accrued in the non-transfused subgroup, but with paradoxically longer survival. Despite different study populations and a span of over 20 years across the seven trials, the reported mortality effects were consistently null, summary RR = 0.99 ($p = 0.864$, homogeneity $p = 0.709$).
Conclusions: HSD delayed blood transfusion by modifying standard triggers like SBP with no detectable effect on survival. The reported heterogeneous HSD effects in subgroups can be explained by collider bias that trauma researchers can avoid by improved covariate selection and data capture strategies.

© 2015 Elsevier Ltd. All rights reserved.

Introduction

Well recognised now, survival bias, or the failure to exclude antecedent survival time that the intervention under study could not possibly affect, has cast doubt on many published reports of trauma comparative effectiveness research [1–7]. While randomisation controls survival and other types of bias [8], insidious collider bias continues to plague trauma trials as well as observational studies. Collider bias is a type of selection bias that is often introduced inadvertently into trauma resuscitation research by restriction, stratification or adjustment on a covariate that is a consequence (collider) rather than a cause (confounder) of both the treatment and outcome of interest. A familiar example of collider bias is Berkson's bias or fallacy [9]. In his 1946 landmark paper [10], renowned physician and statistician, Joseph Berkson, recounted how physicians hypothesised that cholecystic disease caused or aggravated

Report Documentation Page

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 MAY 2015	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Collider bias in trauma comparative effectiveness research: The stratification blues for systematic reviews	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
del Junco, Deborah J. Bulger, Eileen M. Fox, Erin E. Holcomb, John B. Brasel, Karen J. Hoyt, David B. Grady, James J. Duran, Sarah Klotz, Patricia Dubick, Michael A. Wade, E.	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
United States Army Institute of Surgica; Research, JBSA Fort Sam Houston, TX 78234	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**6**	

diabetes after observing a correlation between the two diseases in hospitalised patients. In his hypothetical data analysis, Berkson demonstrated that, although the two diseases were truly not associated with one another in the total population at risk, a spurious positive association arose in an analysis restricted to hospitalised patients. The corresponding analysis for the remaining non-hospitalised population produced an equal and opposite spurious negative association. These puzzling findings were the result of subdividing the total population on an invalid covariate, hospital admission, a consequence (outcome) of both cholecystic disease and diabetes rather than a risk factor or contributing cause (e.g., age or dietary history).

Collider bias is common in trauma resuscitation research due to the difficulty of ascertaining the extent and severity of injury and haemorrhage before initiating the intervention of interest, and the use of poor proxies often not established until hours later. Of the seven trials of prehospital hypertonic saline in dextran (HSD) for hypovolaemic trauma patients [11–17] evaluated in systematic reviews [18–26], none found an overall between-group difference in survival, but four reported significant subgroup effects [11,12,14,16]. Reconciling the promising subgroup results with consistently null overall findings has been challenging [18–26]. We hypothesised that a preventable type of collider bias could explain the incongruities.

Methods

Source data

The two most recent HSD trials, a single-site pilot [11] and a multi-site pivotal study [12], provided data for a secondary analysis to assess collider bias. The trials were funded by the National Heart Lung and Blood Institute of the National Institutes of Health and approved by the U.S. Food and Drug Administration [11,12,27]. The trials compared the effects of prehospital infusion of an initial 250 ml bolus of crystalloid solution containing either HSD (7.5% NaCl with 6% dextran-70) or the standard-of-care, i.e., lactated Ringer's in the pilot [11]; normal saline (0.9%) in the pivotal trial [12]. The trials had similar designs with a few exceptions [11,12,27]. The pivotal trial included a third arm, hypertonic saline without dextran [12]. We excluded patients in that arm ($N = 256$) due to the inability to test for homogeneity across the two trials and hypertonic saline solutions. The pivotal trial [12] ($N = 596$) included patients with penetrating or blunt injuries, and patients with a prehospital systolic blood pressure (SBP) $\geq$70 mmHg were required to have a prehospital heart rate $\geq$108 beats per minute to enrol only hemorrhagic shock patients. The pilot study [11] ($N = 209$) enrolled blunt trauma patients only, required prehospital SBP $\leq$ 90 mmHg, and defined the outcome as 28 day survival free of acute respiratory distress syndrome. De-identified data sets and data dictionaries were acquired and verified. Our Institution's Committee for the Protection of Human Subjects approved the project.

Criteria for a valid covariate

The prevailing guidance for comparative effectiveness research asserts that, regardless of their intended purpose as a confounder or effect modifier, valid covariates or stratification variables must be either defined with known patient status before treatment begins or known to be unaffected by the treatment [28,29]. Variables that either change in response to treatment or develop only post-randomisation are not considered valid covariates. We applied these criteria and constructed causal diagrams to evaluate whether massive transfusion (MT), $\geq$10 units of red blood cell (RBC) transfusions within 24 hours of admission to the emergency department (ED), status fulfilled the criteria for a valid covariate.

Statistical analyses

Following the a priori statistical analysis plan, patient data in the original trials [11,12] were stratified by the post-randomisation covariate and well-established surrogate for bleeding severity, MT [1,3,6,30]. Because randomisation was expected to balance potential confounders across study arms, the data were stratified by MT to reveal HSD modifying effects, or whether the MT subgroup benefited more than the alternate subgroups (receiving 1–9 and 0 RBCs, respectively) [11,12]. Despite favourable HSD effects in the MT subgroup, opposite effects in the 0-RBC subgroup halted the pivotal trial early. To assess whether collider bias could explain null overall, but significant MT subgroup effects, we conducted secondary analyses of standardised data elements combined from the two trials [11,12] and performed a meta-analysis summarising results from all seven trials [11–17]. Univariate and multivariable analyses included chi square tests, the Breslow–Day test for homogeneity of stratum-specific effect estimates, logistic regression and Cox proportional hazards modelling for HSD mortality outcomes, t tests for 24-h RBC counts, linear regression for HSD-induced changes from prehospital to admission SBP, adjusted for prehospital SBP and Wilcoxon rank sum tests for differences in median survival times. Covariates for the mortality analyses included trial identity (pilot [11] or pivotal [12]) and 24-h RBC count. Our meta-analysis of the 1695 patients in 7 trials used the fixed effects confidence interval method [11–17]. Hypothesis tests were two-sided and p values less than 0.05 were considered statistically significant. Data were analyzed with SAS version 9.3 (SAS Institute, Cary, NC) and Stata release 12 (StataCorp, College Station, TX).

Results

To fulfil the criteria for a valid covariate, a hypothetical early indicator of bleeding severity would have been ascertained *before* the infusion of prehospital crystalloid (Fig. 1a). As a valid covariate, the hypothetical early indicator of bleeding severity could have either confounding or modifying effects that appropriate stratification or modelling strategies would reveal [28,29]. However, because resuscitation with hypertonic saline is known to influence the indications for blood transfusion in trauma patients (e.g., increasing SBP) [13–26], and the receipt of one or more RBC transfusions depends on the duration of injury survival (Fig. 1b), the 24 h sum count of RBC transfusions used for stratification in the HSD trials fulfill the definition for an invalid collider covariate [31,32]. Causal diagramming using a directed acyclic graph or DAG (Fig. 1c) [33] shows the irreversible path between prehospital HSD infusion and the 24 h sum of RBCs precluding meaningful interpretation of mortality analyses stratified by RBC category. The true association between HSD and mortality is confounded by the spurious association introduced by the causal effects of early mortality and HSD on the 24 h sums of RBC transfusions, or reverse causation [33]. The well-known peak mortality rates in the minutes to hours following injury occurrence [34,35] limit haemorrhaging patients' opportunity to receive RBC transfusions.

In the combined trial data ($N = 805$) [11,12], overall 30-day mortality was similar in the HSD and control groups (Table 1). Across the arms of the two trials, randomisation appeared to balance potential confounders, however, in the pilot trial, a larger proportion of HSD than control patients were transported by helicopter and intubated [11]. The overall mortality relative risk (RR for HSD vs. control) was somewhat higher (RR = 1.31, 95% CI = 0.82–2.09) than that of the pivotal trial (RR = 1.01, 95% CI = 0.77–1.35), though the two estimates did not differ significantly (p value = 0.363). HSD and control crystalloid had similar effects on mortality (Table 1), across three time intervals following
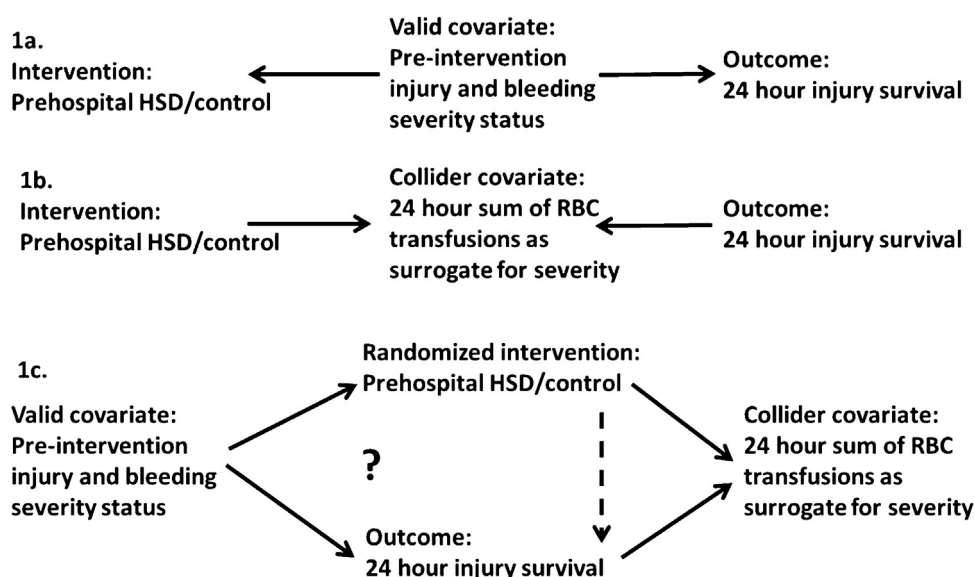
Fig. 1. Causal directions for a valid covariate (a), a collider covariate (b) and their causal pathways in a directed acyclic graph or DAG (c) of prehospital HSD trials. (a) *Valid covariate*: Causal arrows point from valid covariate (pre-intervention injury and bleeding severity status) to both intervention (prehospital HSD or control crystalloid) and outcome (24 h injury survival). (b) *Collider bias*: Causal arrows point from both intervention and outcome to collider (surrogate severity indicator, 24 h sum of RBC transfusions). (c) *DAG*: Solid lines depict causal paths. A valid covariate causes both intervention and outcome. Stratifying on a valid covariate would reveal the true causal path between the intervention and outcome (the "?" represents the unknown causal relationship, if any, that the trial is intended to determine). A collider is a consequence of both intervention and outcome. The dashed line reflects the spurious association between the intervention and outcome introduced by stratification on a collider (invalid) covariate. The direction of the causal arrow at the bottom of the dashed line reflects misinterpretation of the spurious association as a causal effect of treatment on outcome.

enrolment (homogeneity *p* value = 0.339). Among HSD and control decedents, there were no differences in times to death overall or when subdivided into time intervals after enrolment (Table 1). Mortality hazard ratios estimated from Cox proportional hazards models (data not shown) were very similar to the RRs estimated from the logistic regression models. There were no between-group differences in the 24 h sum count of RBCs transfused (Table 1).

As observed in previous trials [13–26], among patients surviving to ED admission, prehospital to admission SBP increased significantly more for HSD than control patients in the combined data set (Table 1). Given that low or declining SBP can be one of the indications for transfusion, we explored the hypothesis that the greater SBP increases among HSD patients (Table 1) reflected HSD-induced physiologic changes that may have delayed the initial RBC transfusion. Because the

trials had not captured data on the timing of RBC transfusions, we sought indirect evidence of delayed initial transfusion in the HSD group using data available on changes in SBP and survival time. The difference between HSD and control group SBP changes (adjusted for prehospital SBP) was greater for decedents than survivors (Table 1) and the differences tended to increase with decreasing survival times (data not shown).

Stratification by the 24-h RBC count confirmed the significantly heterogeneous overall mortality differences across RBC subgroups previously reported for both trials [11,12] (Table 2). The most extreme heterogeneity occurred in the earliest time interval (Table 2). However, patients dying during transport to the ED had no chance for RBC transfusion, and their mortality did not differ between study groups (4/330 = 0.012, HSD, vs. 6/475 = 0.013,

Table 1
Overall between-group differences in outcome (before stratification) using combined data set.

Outcome	HSD group	Control group	Relative risk or difference	*p* Value	95% Confidence interval
Overall 30 day mortality	89/330 = 0.27	118/475 = 0.25	RR = 1.09	0.497	RR = 0.86–1.37
Mortality within 6 h	48/330 = 0.14	73/475 = 0.15	RR = 0.95	0.748	RR = 0.68–1.32
Mortality between 6 and 24 h	12/282 = 0.04	17/402 = 0.04	RR = 1.01	0.986	RR = 0.49–2.07
Mortality between 1 and 30 days	29/270 = 0.11	28/385 = 0.07	RR = 1.48	0.121	RR = 0.90–2.42
Mean no. of days to death overall (SD)	2.10 (4.57)	1.95 (4.33)	Diff = 0.15	0.814	Diff = −1.37–1.08
Mean no. of hours to death between 0 and 6	1.81 (1.17)	2.25 (1.46)	Diff = −0.446	0.080	Diff = −0.05–0.94
Mean no. of hours to death between 6 and 24	6.77 (5.44)	7.21 (6.34)	Diff = −0.44	0.847	Diff = −4.19–5.08
Mean no. of days to death between 1 and 30	5.15 (6.35)	6.65 (6.07)	Diff = −1.50	0.365	Diff = −1.80–4.80
Mean 24 h sum RBC count (SD)	5.03 (8.09)	5.21 (8.74)	Diff = −0.18	0.774	Diff = −1.02–1.37
% In 0 RBC subgroup	136/329 = 41.3%	189/473 = 40.0%	Diff = 1.3%		
% In 1–9 RBC subgroup	133/329 = 40.4%	208/473 = 44.0%	Diff = −3.6%		
% In ≥10 RBC subgroup	60/329 = 18.2%	76/473 = 16.1%	Diff = 2.1%	0.548[*]	
Mean SBP change Δ, prehospital to ED admission —overall	Pre = 64.11 N = 318 Δ = 51.79 (44.06)	Pre = 62.09 N = 458 44.62 (46.37)	Adjusted[†] Diff = 8.61	0.003	Adjusted[†] Diff = 2.93–14.29
Died within 30 days	Pre = 46.35 N = 83 Δ = 44.85 (59.88)	Pre = 52.09 N = 109 27.76 (33.12)	A[†]Diff = 13.52	0.089	A Diff = −2.00–29.04

[*] *p* Value for overall between-group difference in proportionate distribution across 3 RBC subgroups.
[†] Adjusted for pre-hospital SBP. A Diff = adjusted difference.

Table 2
Group mortality and relative risk, by time of death and 24 h sum RBC count.

Outcome	HSD group	Control group	Relative risk	p Value	95% Confidence interval
Overall 30 day mortality				0.023[*]	
0-RBC subgroup	30/136 = 0.22	22/189 = 0.12	1.89	0.012	1.14–3.14
1–9 RBC subgroup	29/133 = 0.22	51/208 = 0.24	0.89	0.564	0.60–1.33
≥10 RBC subgroup	30/60 = 0.50	44/76 = 0.58	0.86	0.359	0.63–1.19
Mortality within 6 h				0.065[*]	
0-RBC subgroup	19/136 = 0.14	17/189 = 0.09	1.55	0.159	0.84–2.88
1–9 RBC subgroup	17/133 = 0.13	28/208 = 0.13	0.95	0.856	0.54–1.67
≥10 RBC subgroup	12/60 = 0.20	27/76 = 0.355	0.56	0.047	0.31–1.01
Mortality between 6–24 h				0.616[*]	
0-RBC subgroup	4/117 = 0.03	4/172 = 0.02	1.47	0.578	0.37–5.76
1–9 RBC subgroup	2/116 = 0.02	6/180 = 0.03	0.52	0.405	0.11–2.52
≥10 RBC subgroup	6/48 = 0.13	7/49 = 0.14	0.88	0.796	0.32–2.41
Mortality between 1 and 30 days				0.085[*]	
0-RBC subgroup	7/113 = 0.06	1/168 = 0.006	10.41	0.006	1.30–83.44
1–9 RBC subgroup	10/114 = 0.09	17/174 = 0.10	0.90	0.776	0.43–1.89
≥10 RBC subgroup	12/42 = 0.29	10/42 = 0.24	1.20	0.620	0.58–2.47

[*] p Value for homogeneity test across the three RBC subgroups.

controls, p = 0.949). The subgroup of patients classified as receiving 0 RBCs included patients who died early, before any RBCs were transfused (23 in the HSD group; 21 in the control group), as well as those who survived 24 h without transfusion (113 in the HSD group; 168 in the control group, Table 2). The 7 prehospital HSD trials [11–17] included in the meta-analysis reported consistently null and non-significant in-hospital mortality effects (HSD vs. control RRs ranged from 0.65 to 1.35, homogeneity test p value = 0.709). The overall summary RR was 0.99 (95% CI = 0.85–1.15, p = 0.864, see online supplementary material).

Fig. 2 is a theoretical timeline to illustrate how stratifying the 24-h mortality data by patients' 24-h RBC count could have introduced collider bias. If deaths actually occurred at the same rate in each study group, but HSD-induced physiologic changes delayed the initial RBC transfusion [18,22], then the cut-off point for HSD decedents' maximum pre-transfusion hours to death would be shifted to the right. The right shift would classify proportionally more HSD than control deaths as pre-transfusion and fewer HSD than control deaths as MT (as observed in Table 2). Paradoxically, HSD decedents in the 0-RBC subgroup would have extended median hours to death relative to controls (Fig. 2). In our analysis of the combined data set, although the median hours to death did not differ between the HSD and control groups overall (p = 0.675), in the 0-RBC subgroup the HSD group had more median hours to death than controls (p = 0.048) as predicted by Fig. 2.

Discussion

The results from our causal diagrams, secondary analysis [11,12] and meta-analysis across seven independent study populations with
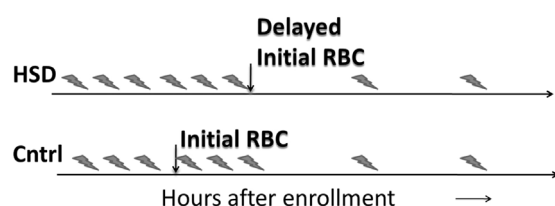

Fig. 2. Theoretical effect of HSD vs. control crystalloid (Cntrl) on delayed transfusion with no effect on survival. If HSD induced physiologic changes and delayed the timing of the initial RBC transfusion (arrows) with no effect on the occurrence or timing of deaths (thunderbolts), then more HSD than control decedents would be classified in the 0-RBC subgroup (left of arrow) and fewer in the alternate subgroups (right of arrow), especially the MT subgroup requiring more time to receive ≥10 transfusions.

different distributions of injury and bleeding severity [11–17] emphasise that, regardless of the HSD-related transfusion delays, deaths among HSD patients occurred at roughly the same rates and points in time as deaths among patients randomised to standard-of-care crystalloid. The protective effect of HSD on survival observed in both trials' MT subgroups was interpreted in the pilot trial report [11] as evidence that prehospital HSD may reduce mortality in the highest risk patients. In contrast, the report of the pivotal trial underscored the adverse HSD effect in the presumed lowest-risk patients receiving 0 RBCs to explain the trial's early termination [12]. Our findings do not support these interpretations, but in no way detract from the decision to terminate the pivotal trial early, given the planned interim trial analyses uncovered potentially adverse HSD effects (e.g., delayed RBC transfusion). They do, however, urge caution in trauma research design and selection of covariates. Assigning patients at the opposite extremes of bleeding severity to the same 0-RBC subgroup reveals how the 24 h RBC classification actually defeats the purpose of stratification, namely, to form more homogeneous patient subgroups [1].

The pivotal trial investigators and expert reviewers considered an HSD-induced delay in initial RBC transfusion as one possible explanation for the adverse findings in the 0-RBC subgroup [12,22]. However, the possibility that HSD could have delayed transfusion without affecting survival, and that the heterogeneous HSD effects on mortality across the different RBC subgroups were artefacts, due solely to bias, was not discussed in the trial reports [11,12], in the accompanying commentary [22] or in subsequent reviews [18,20].

If the suspected HSD-related delay in initial RBC transfusion had truly increased mortality in the HSD trials [11,12,27], the overall analyses would have revealed between-group differences not only in mortality, but also in the time to death and the count of RBC transfusions over 24 h (Table 1). Had there been a true increase in HSD mortality in the subgroup of patients with the lowest bleeding severity (expected to be the largest patient subgroup in either trial), a consistent excess in overall mortality would have been expected in the time period between 1 and 30 days after enrolment. Conversely, had there been a true decrease in HSD mortality in the subgroup of trial patients with the worst bleeding severity (contributing the majority of deaths and in the earliest time period after injury), a consistent deficit in overall mortality would have been expected in the enrolment to 6 h time period. Such time-interval specific and opposite patterns of overall mortality were not reported in the original trials [11,12] or evident from our secondary analyses. Primary cause of death was not captured in the pivotal trial. In the pilot trial, over 53% of the

deaths within 24 h were attributed to hypovolaemic shock. In the time interval between 1 and 30 days, 60% of deaths were attributed to brain injury, but none to hypovolaemic shock.

Because no single, universally accepted indicator of the amount of blood loss or bleeding severity exists for injured patients [36] MT has become entrenched in the trauma research literature as a surrogate marker for the highest severity [1,3,30,37,38,39]. Alternatives have been proposed, for example, the "critical administration threshold" or CAT, defined as a rate of 3 RBC unit transfusions per hour anytime within the first 24 h after emergency medical service activation [30]. Like MT, the CAT surrogate covariate also sums blood product transfusions over time periods during which study patients may die before reaching a rate of 3 RBC transfusions per hour within 24 h. Thus the use of the CAT alternative to MT for restriction, stratification or covariate adjustment is also likely to introduce collider bias.

Like the recent HSD trials [11,12], two previous HSD trials [14,16] reported statistically significant effects in a subgroup defined by a post-randomisation collider covariate (patients undergoing surgery [14] or assigned a score >4 on the Abbreviated Injury Score for the head [16]). Clearly, the potential for collider bias in trauma research is not limited to subgrouping based on the conventional definition of MT or to observational study designs. Distinguishing between a valid and invalid covariate can be difficult because a collider covariate can have a dual role as both a consequence and contributing cause of either the treatment or the outcome. On the one hand, trauma patients must survive long enough to undergo the diagnostic procedures or interventions like MT that may be used as surrogate markers of severity, and on the other, the surrogate markers may also affect survival. Unfortunately, ambiguity in the causal directions linking covariates like MT with treatment or outcome does not preclude collider bias. The accurate and complete ascertainment of injury severity is an evolving process for many severely injured patients. The process can entail a series of definitive diagnostic procedures based on technologies available only in appropriately equipped centres. Clinical correlates, markers and surrogates of injury severity, like massive transfusion (MT), injury severity score and surgical intervention, are easily ascertained by retrospective review of patient records often without regard for their critical timing and sequence relative to the high but rapidly declining rates of early trauma death. These surrogates are considered "validated" by their strong associations with injury mortality, and have become standard covariates for stratification or adjustment by injury severity, depite the inability of the measures of association to distinguish forward from reverse causation. Because of time-honoured conventions, trusted validation, lack of reliable alternatives, and the subtlety of collider bias, the widespread research use of these surrogate biomarkers of injury severity is likely to persist and frustrate systematic reviews [40,41] due to the unpredictable occurrence and variable impact of collider bias (e.g., conflicting findings) on trauma resuscitation research.

We acknowledge that there may be alternative explanations for the observed null overall HSD effects on mortality in the face of counterbalanced opposite effects across strata defined by the 24-h sum of RBC transfusions [11,12]. However, we believe the available evidence and clinical plausibility support our hypothesis of a survival-neutral HSD-induced delay in initiating transfusion. We interpret the meta-analysis results of all seven prehospital HSD trials as evidence that patients resuscitated with prehospital HSD survived unanticipated transfusion delays [11–17] to mirror the survival of their control counterparts. What remains unanswered is whether patients receiving prehospital standard-of-care crystalloid would have survived similar transfusion delays without the physiologic changes induced by HSD [23]. To determine whether prehospital HSD is superior or equivalent [23] to standard-of-care crystalloid in promoting trauma patients' survival beyond the highest-risk interval and long enough to ensure definitive diagnosis and interventions, the following study conditions would be required: (1) a trauma setting in which prolonged emergency transport or unavoidable delays to definitive diagnosis and therapy are the norm, (2) a protocol with reliable indications for ED interventions that (unlike SBP) remain detectable regardless of prehospital fluid resuscitation, and (3) sufficient resources to capture accurate data on the nature and timing of all relevant injury characteristics, interventions and outcomes.

Extreme variation in the extent, severity and duration of serious injury sets off highly dynamic, complex and idiosyncratic sequences of trauma resuscitation interventions. The surge of diagnostic and therapeutic procedures occurs over a compressed timeline with the greatest threat to patient survival immediately following injury [34] and declining precipitously a few hours thereafter [34,35,42]. Thus, some of the challenges to the scientific integrity of trauma comparative effectiveness research may be unique. Standard clinical research training and tools for data analysis are so attuned to the typical increase in patients' mortality rates with the passage of time [1] that even trauma's strikingly opposite pattern can fail to attract attention. Compounding these challenges [39,43] has been the resistance of imminent early trauma death [34] to increasingly swift and sophisticated systems of care [44–46].

The validity and reproducibility of future trauma comparative effectiveness research depends on well-informed leadership and may benefit from (1) explicit a priori diagramming of the credible causal pathways among the key covariates and determinants of outcome, (2) avoiding collider bias by replacing any wrongly-ordered covariates with more logical substitutes, and (3) planning data capture and analysis strategies to account for the timing and sequence of key events and covariates.

Conflict of interest statement

The authors have declared no conflicts.

Funding

Acknowledgment

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.injury.2015.01.043.

References

[1] del Junco DJ, Fox EE, Camp EA, Rahbar MH, Holcomb JB, Group PS. Seven deadly sins in trauma outcomes research: an epidemiologic post mortem for major causes of bias. J Trauma Acute Care Surg 2013;75:S97–103.

[2] Hallet J, Lauzier F, Mailloux O, Trottier V, Archambault P, Zarychanski R, et al. The use of higher platelet: RBC transfusion ratio in the acute phase of trauma resuscitation: a systematic review. Crit Care Med 2013;41:2800–11.

[3] Ho AM, Dion PW, Holcomb JB, Wong RH, Ng CS, Kamakar MK, et al. Reverse survivor bias in observational studies involving cohorts: a lesson from '1:1' trauma studies. Hong Kong Med J 2013;19:461–3.

[4] Ho AM, Dion PW, Yeung JH, Holcomb JB, Critchley LA, Ng CS, et al. Prevalence of survivor bias in observational studies on fresh frozen plasma:erythrocyte ratios in trauma requiring massive transfusion. Anesthesiology 2012;116: 716–28.

[5] Ho AM, Dion PW, Yeung JH, Joynt GM, Lee A, Ng CS, et al. Simulation of survivorship bias in observational studies on plasma to red blood cell ratios in massive transfusion for trauma. Br J Surg 2012;99(Suppl 1):132–9.

[6] Rahbar E, Fox EE, del Junco DJ, Harvin JA, Holcomb JB, Wade CE, et al. Early resuscitation intensity as a surrogate for bleeding severity and early mortality in the PROMMTT study. J Trauma Acute Care Surg 2013;75:S16–23.

[7] Snyder CW, Weinberg JA, McGwin Jr G, Melton SM, George RL, Reiff DA, et al. The relationship of blood product ratio to mortality: survival benefit or survival bias? J Trauma 2009;66:358–62 (Discussion 62-4).

[8] Sturmer T. Study design considerations. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, editors. Developing a protocol for observational comparative effectiveness research: a user's guide. Rockville, MD: Agency for Healthcare Research and Quality; 2013. p. 21–34.

[9] Westreich D. Berkson's bias, selection bias, and missing data. Epidemiology 2012;23:159–64.

[10] Berkson J. Limitations of the application of fourfold table analysis to hospital data. Biometrics 1946;2:47–53.

[11] Bulger EM, Jurkovich GJ, Nathens AB, Copass MK, Hanson S, Cooper C, et al. Hypertonic resuscitation of hypovolemic shock after blunt trauma: a randomized controlled trial. Arch Surg 2008;143:139–48 (discussion 49).

[12] Bulger EM, May S, Kerby JD, Emerson S, Stiell IG, Schreiber MA, et al. Out-of-hospital hypertonic resuscitation after traumatic hypovolemic shock: a randomized, placebo controlled trial. Ann Surg 2011;253:431–41.

[13] Maningas PA, Mattox KL, Pepe PE, Jones RL, Feliciano DV, Burch JM. Hypertonic saline–dextran solutions for the prehospital management of traumatic hypotension. Am J Surg 1989;157:528–33 (discussion 33-4).

[14] Mattox KL, Maningas PA, Moore EE, Mateer JR, Marx JA, Aprahamian C, et al. Prehospital hypertonic saline/dextran infusion for post-traumatic hypotension. The U.S.A. Multicenter Trial. Ann Surg 1991;213:482–91.

[15] Vassar MJ, Fischer RP, O'Brien PE, Bachulis BL, Chambers JA, Hoyt DB, et al. A multicenter trial for resuscitation of injured patients with 7.5% sodium chloride. The effect of added dextran 70. The Multicenter Group for the Study of Hypertonic Saline in Trauma Patients. Arch Surg 1993;128:1003–11 (Discussion 11-3).

[16] Vassar MJ, Perry CA, Gannaway WL, Holcroft JW. 7.5% Sodium chloride/dextran for resuscitation of trauma patients undergoing helicopter transport. Arch Surg 1991;126:1065–72.

[17] Vassar MJ, Perry CA, Holcroft JW. Prehospital resuscitation of hypotensive trauma patients with 7.5% NaCl versus 7.5% NaCl with added dextran: a controlled trial. J Trauma 1993;34:622–32 (Discussion 32-3).

[18] Bulger EM, Hoyt DB. Hypertonic resuscitation after severe injury: is it of benefit? Adv Surg 2012;46:73–85.

[19] Dubick MA, Bruttig SP, Wade CE. Issues of concern regarding the use of hypertonic/hyperoncotic fluid resuscitation of hemorrhagic hypotension. Shock 2006;25:321–8.

[20] Dubick MA, Shek P, Wade CE. ROC trials update on prehospital hypertonic saline resuscitation in the aftermath of the US–Canadian trials. Clinics (Sao Paulo) 2013;68:883–6.

[21] Dubick MA, Wade CE. A review of the efficacy and safety of 7.5% NaCl/6% dextran 70 in experimental animals and in humans. J Trauma 1994;36:323–30.

[22] Holcroft JW. The hypertonic saline trial: a possible downside to the gold standard of double blinding. Ann Surg 2011;253:442–3.

[23] Rocha-e-Silva M. Hypertonic saline for treatment of shock: have we looked for everything? MedicalExpress 2014;1:14–21.

[24] Wade CE, Grady JJ, Kramer GC. Efficacy of hypertonic saline dextran fluid resuscitation for patients with hypotension from penetrating trauma. J Trauma 2003;54:S144–8.

[25] Wade CE, Grady JJ, Kramer GC, Younes RN, Gehlsen K, Holcroft JW. Individual patient cohort analysis of the efficacy of hypertonic saline/dextran in patients with traumatic brain injury and hypotension. J Trauma 1997;42:S61–5.

[26] Wade CE, Kramer GC, Grady JJ, Fabian TC, Younes RN. Efficacy of hypertonic 7.5% saline and 6% dextran-70 in treating trauma: a meta-analysis of controlled clinical studies. Surgery 1997;122:609–16.

[27] Brasel KJ, Bulger E, Cook AJ, Morrison LJ, Newgard CD, Tisherman SA, et al. Hypertonic resuscitation: design and implementation of a prehospital intervention trial. J Am Coll Surg 2008;206:220–32.

[28] Sauer B, Brookhart MA, Roy JA, VanderWeele TJ. Covariate selection. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, editors. Developing a protocol for observational comparative effectiveness research: a user's guide. Rockville, MD: Agency for Healthcare Research and Quality; 2013. p. 93–108.

[29] Varadhan R, Seeger J. Estimation and reporting of heterogeneity of treatment effects. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, editors. developing a protocol for observational comparative effectiveness research: a user's guide. Rockville, MD: Agency for Healthcare Research and Quality; 2013 . p. 35–44.

[30] Savage SA, Zarzaur BL, Croce MA, Fabian TC. Redefining massive transfusion when every second counts. J Trauma Acute Care Surg 2013;74:396–400 (Discussion-2).

[31] Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. Epidemiology 2003;14:300–6.

[32] Janszky I, Ahlbom A, Svensson AC. The Janus face of statistical adjustment: confounders versus colliders. Eur J Epidemiol 2010;25:361–3.

[33] Sauer B, VanderWeele TJ. Use of directed acyclic graphs. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, editors. Developing a protocol for observational comparative effectiveness research: a user's guide. Rockville, MD: Agency for Healthcare Research and Quality; 2013. p. 177–84.

[34] Sobrino J, Shafi S. Timing and causes of death after injuries. Proc (Bayl Univ Med Cent) 2013;26:120–3.

[35] Demetriades D, Murray J, Charalambides K, Alo K, Velmahos G, Rhee P, et al. Trauma fatalities: time and location of hospital deaths. J Am Coll Surg 2004;198:20–6.

[36] Brockamp T, Nienaber U, Mutschler M, Wafaisade A, Peiniger S, Lefering R, et al. Predicting on-going hemorrhage and transfusion requirement after severe trauma: a validation of six scoring systems and algorithms on the TraumaRegister DGU®. Crit Care 2012;16:R129.

[37] Mitra B, Cameron PA, Gruen RL, Mori A, Fitzgerald M, Street A. The definition of massive transfusion in trauma: a critical variable in examining evidence for resuscitation. Eur J Emerg Med 2011;18:137–42.

[38] Sharpe JP, Weinberg JA, Magnotti LJ, Maclennan PA, Schroeppel TJ, Fabian TC, et al. Accounting for differences in transfusion volume: are all massive transfusions created equal? J Trauma Acute Care Surg 2012;72:1536–40.

[39] Holcomb JB, del Junco DJ, Fox EE, Wade CE, Cohen MJ, Schreiber MA, et al. The prospective, observational, multicenter, major trauma transfusion (PROMMTT) study: comparative effectiveness of a time-varying treatment with competing risks. JAMA Surg 2013;148:127–36.

[40] Bhangu A, Nepogodiev d, Doughty H, Bowley DM. Meta-analysis of plasma to red blood cell ratios and mortality in massive blood transfusions for trauma. Injury 2013;44:1693–9.

[41] Hallet J, Lauzier F, Mailloux O, Trottier V, Archambault P, Zarychanski R, Turgeon AF. The use of higher platelet:RBC transfusion ratio in the acute phase of trauma resuscitation: a systematic review. Crit Care Med 2013;41:2800–11.

[42] Trunkey DD. Trauma accidental and intentional injuries account for more years of life lost in the U.S. than cancer and heart disease. Among the prescribed remedies are improved preventive efforts, speedier surgery and further research. Sci Am 1983;249:28–35.

[43] Arbogast PG, VanderWeele TJ. Considerations for statistical analysis. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, editors. Developing a protocol for observational comparative effectiveness research: a user's guide. Rockville, MD: Agency for Healthcare Research and Quality; 2013 . p. 135–44.

[44] Bailey J, Trexler S, Murdock A, Hoyt D. Verification and regionalization of trauma systems: the impact of these efforts on trauma care in the United States. Surg Clin North Am 2012;92:1009–24. ix–x.

[45] Hoyt DB, Coimbra R. Trauma systems. Surg Clin North Am 2007;87:21–35. v–vi.

[46] MacKenzie EJ, Rivara FP, Jurkovich GJ, Nathens AB, Frey KP, Egleston BL, et al. A national evaluation of the effect of trauma-center care on mortality. N Engl J Med 2006;354:366–78.